ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Procedia CIRP 41 (2016) 746 – 751
www.elsevier.com/locate/procedia

48th CIRP Conference on MANUFACTURING SYSTEMS - CIRP CMS 2015

Joint Angle Data Representation for Data Driven Human Motion Synthesis

Han Du[a]*, Martin Manns[b], Erik Herrmann[a], Klaus Fischer[a]

[a] DFKI GmbH, Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany
[b] Daimler AG, Wilhelm-Runge-Str. 11, 89081 Ulm, Germany

* Corresponding author. Tel.: +49-681-85775-5395. *E-mail address:* han.du@dfki.de

Abstract

For ergonomic assessment of manual assembly tasks, digital simulation has received increasing attention due to its efficiency compared to physical prototypes. One of the crucial parts of digital simulation is an accurate animation of the digital human model (DHM). Current digital simulation tools such as Delmia V5 require interactive manual editing to produce animations, which is time consuming and can look unnatural. On the other hand, data-driven motion synthesis that is based on motion capture data can produce natural motions with little user involvement. The practical difficulty lies in processing motion data into a parameterized motion model. A common approach is decomposing motions and categorizing them into finite short motion primitives. For each motion primitive, motion data is represented as a numerical vector, on which functional principal component analysis (FPCA) is applied to reduce dimensionality. In this work, different ways of representing joint angles from motion capture data are explored: Euler angle, quaternion and exponential map. The data representations are evaluated for their reconstruction error with FPCA. In the tests, quaternion representation shows best performance for motion data representation, which contradicts a preference in literature for exponential map representation. Therefore, quaternion representation is considered appealing for statistically modelling motion data.

Peer-review under responsibility of the scientific committee of 48th CIRP Conference on MANUFACTURING SYSTEMS - CIRP CMS 2015

Keywords: Assembly; Simulation; Functional principal component analysis

1. Introduction

In large series production such as in automotive industry, manual assembly processes are common – especially for variant rich products. These processes are thoroughly planned and verified so that processes on the shop floor are stable and efficient at start of production (c. [1]). One essential point of verification is ergonomics, which ensures that workers are able to carry out assembly activities for long periods without negative health impact (s. [2]).

The traditional way of ergonomic assessment for manual assembly process requires a physical prototype [3], i.e. an early release of the product [4]. With the ongoing trend to shorter life cycles and increased numbers of product variants [5], more prototypes are required for process verification.

Simulation of manual assembly processes has the potential of reducing prototype cost (c. [5]). For example, Delmia V5 is used at Daimler for manufacturing design and verification. While Delmia V5 provides high quality simulations, modeling requires manually specifying poses for the virtual human model. The simulation of a task of one minute simulated time can require several hours of modeling time [7].

Digital simulation tools such as EMA that reduce modeling effort are available [8]. However, naturalness of motions is visibly reduced so that some motions are hard to identify (c. [7]), which reduces ergonomics experts' acceptance.

One popular approach to address naturalness is motion capture data driven motion synthesis. Data driven motion synthesis approaches that derive new motions from previously recorded motion capture data promise high motion quality while keeping modeling effort low. However, data driven motion synthesis suffers from the high-dimensional input data sets that result from motion capture. Therefore, dimension reduction is required. A method that is widely employed is functional principal component analysis (FPCA, c. e.g. [9]).

Peer-review under responsibility of the scientific committee of 48th CIRP Conference on MANUFACTURING SYSTEMS - CIRP CMS 2015
doi:10.1016/j.procir.2015.12.096

For FPCA to effectively reduce input dimensions, adequate data representation is crucial.

This work focuses on skeleton based motion capture data with fixed bone lengths. Therefore, input data consists of a root node position that is accompanied by a set of joint angles.

There are several publications that report representing these angles in different ways. Tilmanne and Dutoit suggest on a principal basis that the exponential map representation should be employed [10].

In the following sections, three joint angle representations are compared for the motion types walk, carry, pick and place: Euler angles, quaternions and exponential map. This analysis derives experimental evidence, which of the representations is suited best for FPCA based dimension reduction.

2. Motion Data Parameterization Methods

The task of motion data parameterization is to numerically represent motion capture data for further data analysis. Many motion capture systems track joint markers of human skeletons [11]. An intuitive way to represent motion data is to directly use the global 3D Cartesian position of each joint of the skeleton. Then the full-body pose is represented as a vector of ordered joints position. However, this approach is not widely used in statistical analysis of motion data, because the skeleton is not implicitly preserved, and new variances introduced by statistical methods could break the skeleton during virtual human action.

This work focuses on motion data in a relative translation and orientation parameterization method. This type of methods is widely used in statistical human motion modeling [12][13][14]. The full-body pose is represented as root position and orientation as well as joint relative orientation. One of the main appealing properties to use this parameterization is that skeleton information is implicitly preserved. However, parameterizing three degree-of-freedom (DOF) rotation is a nontrivial problem.

Euler angles are widely used to represent joint orientation [15]. An Euler angle defines a rotation about one of the coordinate axes. So for three DOF, three Euler angles are used for three single-axis rotations. Euler angle representation is easy to understand the meaning of each variable. However, three DOF Euler rotations suffer from gimbal lock [16], which means one DOF is “locked up”. Furthermore, Euler angles represent 3D rotation as rotations about x, y, z axes independently, ignoring the correlation between x, y, z rotations. Another problem about Euler angles are singularities [15], which means that two set of rotation angles may correspond to the same rotation.

Quaternions have a long history for representing rotation in computer graphics [16]. Quaternions are hypercomplex values with a real part and three imaginary parts that represent a rotation in three degrees of freedom. A quaternion is defined as follows:

$$q = [q_w, q_x, q_y, q_z] \tag{1}$$

where q is quaternion vector, q_w is the real part of quaternion, q_x, q_y, q_z are imaginary part of quaternion.

Since unit quaternions are free from gimbal lock, quaternion must be normalized to achieve this property. Furthermore, quaternions suffer from the singularity problem, e.g. q and $-q'$ correspond to the same rotation.

The Exponential map is in general a re-parameterization of a quaternion [17]. It maps vectors in R^3 into unit quaternion:

$$\boldsymbol{q} = e^{v} = [\sin(\frac{1}{2}\theta)\frac{\boldsymbol{v}}{\theta},\ \cos(\frac{1}{2}\theta)] \tag{2}$$

where $\boldsymbol{v} \in R^3, \|\boldsymbol{v}\| = \theta$.

Exponential map has the benefit that it linearizes quaternions [15]. However, since an exponential map is a conversion of a quaternion, the singularity problem still exists.

3. Motion Data Preprocessing

Generally, data driven motion synthesis is not conducted on raw but on preprocessed motion capture data. These preprocessing steps remove inner structure and redistribute variance across values. For the experimental setup of this work, structurally similar motions are decomposed into motion primitives (e.g. a walk motion into right stance and left stance). Next, these cut motions are aligned against each other (c. [18]).

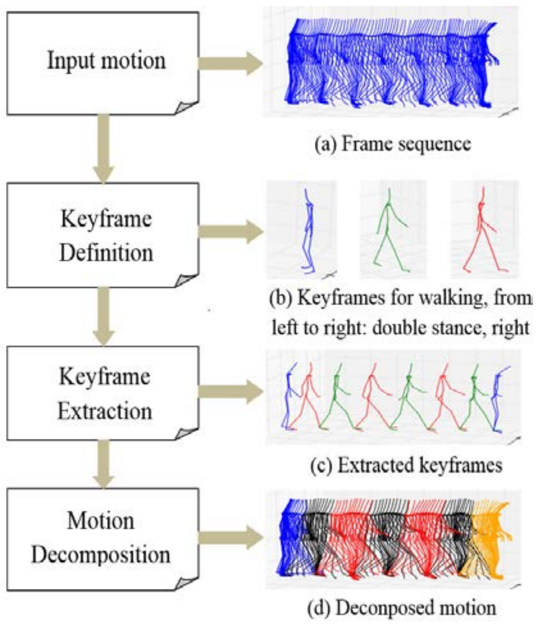

Fig. 1. Work-flow of motion decomposition: (a) the input frame sequence; (b) defined key frames, from left to right: double stance, right stance and left stance; (c) automatically extracted key frames from frame sequence; (d) segment motion into clips and categorize them into different motion primitives based on key frames.

3.1. Motion Decomposition

Although each human action displays a large range of variations, the high level structures can be considered finite [18]. Fig. 1.a shows a sequence of walking. Some patterns are repeated as a sequence of alternating left step, right step and double step. These patterns can be defined by labelling key frames in the frame sequence and segment motions into small clips that are separated by key frames. For instance for normal walking, three key poses are defined: Double stance, left stance and right stance (see Fig. 1.b). If clips share the same starting and ending key frames, they are categorized into the same motion primitive. Our knowledge about motion style is also embedded into each motion primitive, because the clips in one motion primitive share some common properties. For example, the contact information is always consistent for walking.

A good decomposition should make the similarity large for motion segments within the same motion primitive, and small for motion segments in different motion primitives. Based on this criterion, we define two kinds of frame as key frames:

- Frame instances when a contact state transition occurs. For example, during walking, foot contact with ground alternatively changes (e.g. left foot strike).
- Frame instance when highest visual content changes.

For walking, the extraction of key frames is detected by extracting dominant features such as contact points in the frame sequence (see Fig. 1.c). For picking and placing, key frame labelling is done manually in order to avoid errors.

3.2. Motion Alignment

After motion decomposition, motion clips in the same motion primitive may have different length, orientation and position. A translation and rotation invariant frame distance (following [19]) is applied to measure the similarity between frames, and dynamic time warping techniques (following [20]) are used use to align all motion clips in one motion primitive to the reference motion clip. The reference motion clip is chosen by finding the one with minimal average frame distance to other clips in one motion primitive. The canonical timeline $t = 1, ..., T$ is defined by reference motion clip, and all the other clips with different timeline $t = 1, ..., T'$ are mapped to this canonical timeline. After motion alignment, each motion clip $M(t)$ is decoupled as geometric parameter $s(t)$ and temporal parameter $w(t)$.

3.3. Motion Data Smoothing

For testing quaternion representations in a practical relevant setting, the singularity problem is addressed by applying smoothing on quaternions. Fig. 2 (top) displays quaternion values for one joint over frames extracted from data after post-processing. It is visible that sign flipping occurs frequently, which comes from motion alignment. This phenomenon breaks the assumption of FPCA based dimension reduction that motion data changes smoothly over time.

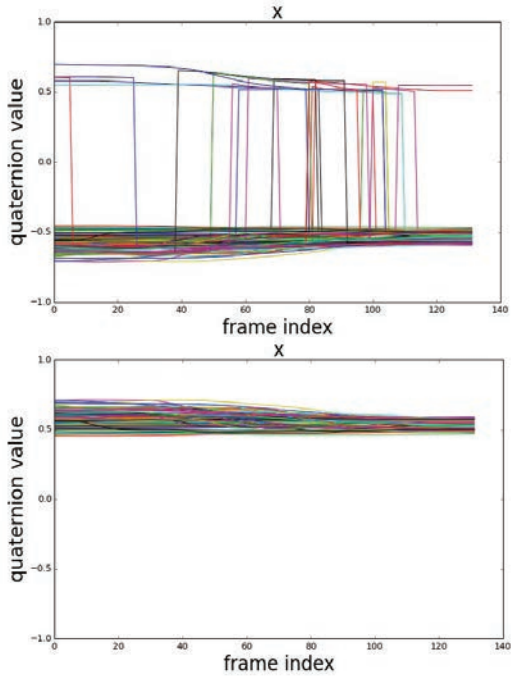

Fig. 2. Quaternion component for the right shoulder before (top) and after (bottom) smoothing.

In order to eliminate this singularity problem, for each frame, the closest matching quaternion for the next frame is chosen, where angular distance between quaternion is considered:

$$dist(\boldsymbol{q}, \boldsymbol{q}') < dist(\boldsymbol{q}, -\boldsymbol{q}') \quad if \;\; dist(\boldsymbol{q}, \boldsymbol{q}') > 0 \tag{3}$$

where q and $-q'$ are two quaternions having opposite orientation.

4. Testing Angle Representations for Dimension Reduction

4.1. Angle representation

Each aligned motion clip $s(t)$ in a motion primitive is represented as a numerical vector with equal length to each other by sequentially concatenating pose parameters of each frame after motion data pre-processing.

For Euler angle representation, the vector starts with the Cartesian root joint position, which is followed by the first frame’s Euler angles of each joint, then the second frame’s and so forth. For quaternion and exponential map representations, Euler angles are replaced with the respective values. The resulting raw motion data has more than 10000 dimensions for each aligned motion clip.

Before applying dimension reduction, all motion vectors are normalized because root joint position and angle representation are not in the same unit. In order to avoid influence of larger position vector elements of the root node to overall variance, the root positions are normalized into the same range as each angle representation.

4.2. Functional PCA

Functional PCA is a variant of standard PCA, which regards multivariate data as samples from a continuous function. There are different implementations of Functional PCA [21]. In this work, it is assumed that joint orientation parameters are continuously changing over frames. Therefore, joint angle trajectories can be interpolated with a smooth function, which is defined as an expansion of basis functions:

$$\{x_{i1},...,x_{in}\} \rightarrow x_i(t) \rightarrow \sum_{k=1}^{K} c_{ik}\phi_k(t), \quad K << n \tag{4}$$

where x_{in} denotes quaternion value of i-th joint in n-th frame, $x_i(t)$ denotes a continuous function, $\phi_k(t)$ denotes cubic b-spline basis functions.

Each function $x_i(t)$ is represented by coefficients of the basis functions. So for each joint parameter, the dimension is reduced from n to K. Each motion is represented as a vector of sequentially concatenated coefficients of each joint parameter. Standard principal component analysis is applied to the functional representation of motion data. Therefore, motion clips can be represented the mean motion $\boldsymbol{p}_0$ and weighted sum of Eigenvectors $\boldsymbol{p}_i$:

$$\boldsymbol{s} = \boldsymbol{p}_0 + [\boldsymbol{p}_1,..., \boldsymbol{p}_M]\boldsymbol{\alpha} \tag{5}$$

where $\boldsymbol{s}$ is functional representation of motion data, $\boldsymbol{p}_0$ denotes the mean motion, $\boldsymbol{p}_i$ denotes eigenvector of functional data, $\boldsymbol{\alpha}$ denotes the eigen-weights vector.

5. Experimental Results

5.1. Motion Capture Data

The motion capture data used for the experiments is captured using an Optitrac Flex13 system with three different actors. The captures motion data set contains four elementary actions: walking, carrying, picking and placing performed in assembly workshop scenario. The details of the data that is used can be found in Table 1.

Table 1. Motion Capture Data.

	walking	carrying	picking	placing
No. captured motions	124	144	194	186
No. primitives per motion	3	3	3	3
No. key frames	6	6	2	2

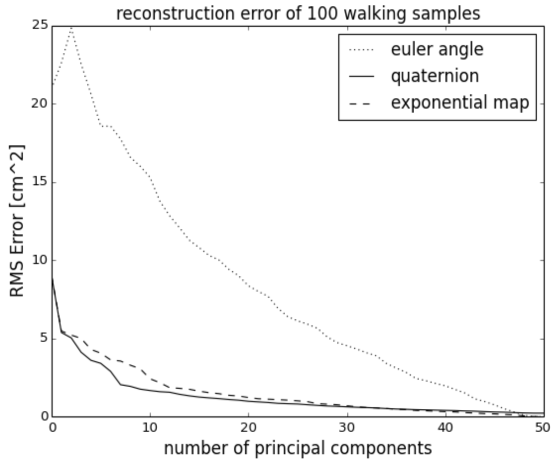

Fig. 3. Reconstruction error of 100 training samples for walk using Euler, quaternion and exponential map representations.

5.2. PCA efficiency

Preprocessing steps described in section 3 are applied to the elementary actions. Table 1 shows how many key frames and motion primitives are found in each elementary action.

For each motion sample, the root mean square (RMS) error is measured between the reconstructed motion sample from dimension reduction and original motion sample. It is computed by averaging the error of global joint positions between the original motion and the reconstructed motion. Therefore, the RMS is independent from the angle parameterization methods.

Fig. 3 shows the root mean square (RMS) error of full body of reconstructing 100 training samples of picking from functional PCA. Euler angle shows the highest RMS errors while quaternion shows lowest. RMS errors for exponential map representation are higher than for quaternion representation for low numbers of principal components. For numbers larger than 30 quaternion and exponential map representations yield comparable results.

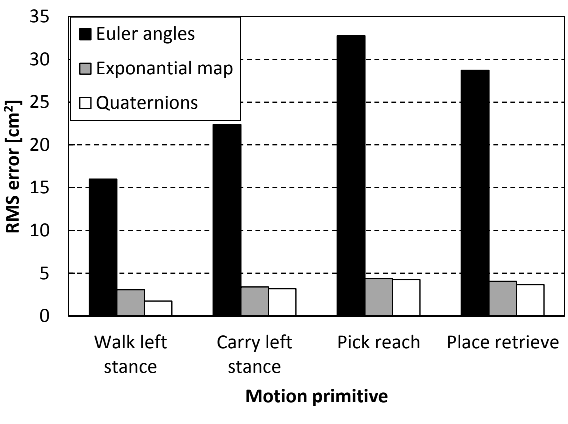

Fig. 4. Reconstruction error for four actions using ten principal components

Note that no smoothing is applied for Euler angles and that without smoothing, Euler angles still yields higher RMS error then quaternion or exponential map representations.

Fig. 4 shows reconstruction errors for all parameterization methods of four elementary actions using first 10 principal components. The number has been chosen because this number is considered adequate for data driven motion synthesis.

RMS errors are lower for walking and carrying, and higher for picking and placing. Quaternion representation shows a visibly lower RMS error for walking than exponential map. For carrying, picking and placing, the order of RMS error stays the same. However, the RMS error difference is smaller.

6. Discussion

While Johnson [17] suggests that exponential map is preferable for statistical modeling model if linear dimension reduction approaches are employed our empirical analysis shows that quaternions may be betters suited.

This benefit could be reproduced for different types of motions. In general, reconstruction error increases with the increase of the complexity of motions. From Fig. 4, it is visible that Euler angle representation results are quite sensitive to the complexity of motion. The benefit of quaternions seems to be largest for simple motions such as walking if the number of principal components is low. For complex motions such as picking and placing, the difference becomes less obvious. However, even for complex motions with many principal components no clear benefit of exponential map is apparent.

This result is particularly interesting because quaternion representation increases the length of the input vector by almost 4/3 because it requires 4 instead of 3 values that the other representations employ per joint.

Smoothness of data seems more important for FPCA efficiency in our approach than linearity of data. Therefore, Euler angle representation is considered unsuitable data driven motion synthesis that follows the considered FPCA based methodology for dimension reduction. The importance of smoothness is underlined because exponential mal data without smoothing quaternions showed results that are only slightly better than Euler angle representation.

All the motion capture data used in our experiments come from the same motion capture system, and post-processing steps are the same for all data. Therefore, data is considered comparable regarding quality for the different motions. While quaternion based joint angle representation has been successfully applied for other motion capture systems, generalizability of results for different motion capture systems has not been tested.

While the tests employ RMS error as quality measure, it shall be noted that low reconstruction error does not necessary indicate good motion for ergonomics simulation. If e.g. one joint is broken in one frame of a motion then the motion is normally considered unsuitable for simulation. A similar RMS error may be measured for a reconstructed motion that has tiny joint angle errors in each frame and therefore is considered acceptable. However, visual analysis of reconstructed motions makes obvious that at an error level of larger than 10 cm², motion quality is normally considered inacceptable for ergonomics analysis.

Smoothing on Euler angle representation has not been conducted in the tests because for 3 DoF, Euler angle rotate about each axis separately. Smoothing each rotation independently is likely to change the implicit artificial variance of 3D rotation, which would mean an additional, unrelated source of artificial variance for FPCA.

The considered elementary actions walk, carry, pick and place do not represent the majority of actions that are relevant for manual assembly operations. Instead, they are considered well understood, easy to model and exhibit less motion variance than e.g. turning steps. Therefore, the presented results may not be generic for all action types.

7. Conclusion and Outlook

Three different joint angle representations have been applied to a data driven motion synthesis approach that models the elementary actions walk, carry, pick and place. These actions are practically relevant for manual assembly tasks.

Quaternion based motion representation is considered more beneficial to FPCA efficiency than the exponential map representation, while Euler representation is considered unsuitable for data driven motion synthesis. The benefit decreases for complex motions and large number of principal components.

The diminished RMS errors between quaternion and exponential map representations make it hard to extrapolate results for motions that are more complex or variant rich than the four motions considered in the tests. Therefore in future research, the experiments should be extended to more complex motion types such as tool handling and single stances for turning around and side stepping.

Acknowledgements

This work is funded by EU project INTERACT - Interactive Manual Assembly Operations for the Human-Centered Workplaces of the Future that has received funding from the European Union Seventh Framework Programme (FP7, Grant Agreement No. 61107).

The authors would like to thank the Institute of Media Informatics at Ulm University, Germany for providing motion capture data and the Laboratory for Manufacturing Systems & Automation at the University of Patras, Greece for their help preprocessing the motion capture data.

References

[1] Bracht U, Geckler D, Wenzel S. Digitale Fabrik. Springer-Verlag Berlin Heidelberg. 2001.
[2] Fritzsche L, Wegge J, Schmauder M, Kliegel M, Schmidt K. Good ergonomics and team diversity reduce absenteeism and errors in car manufacturing. Ergonomics. 2014; 57 (2), p. 148–161.
[3] Nof S, Wilhelm W, Warnecke H, Industrial Assembly, London, Weinheim, New York, Tokyo, Melbourne, Madras. 1997.
[4] Hartmann B. Reflective physical prototyping through integrated design, test, and analysis. In Proceedings of UIST. 2006; p. 299–308.

[5] Manns M, Arteaga M. A Vagueness Measure for Concurrent Engineering of Manual Assembly Lines. In: Enabling Manufacturing Competitiveness and Economic Sustainability. 2014; p. 131-135.
[6] Kulkarni A, Kapoor A, Iyer M, Kosse V. Virtual prototyping used as validation tool in automotive design. Proceedings of the 19th International Congress on Modelling and Simulation, Perth, Australia. 2011; p. 419-425,.
[7] Manns M, Arteaga M. Automated DHM Modeling for Integrated Alpha-Numeric and Geometric Assembly Planning. Proceedings of the 23rd CIRP Design Conference, Bochum, Germany, Springer-Verlag Berlin Heidelberg. 2011.
[8] Fritzsche L, Jendrusch R, Leidholdt W, Bauer S, Jäckel T, Pirger A. Introducing ema (Editor for Manual Work Activities) – A New Tool for Enhancing Accuracy and Efficiency of Human Simulations in Digital Production Planning. In: V.G. Duffy (Ed.), Digital Human Modeling. HCII 2011; LNCS 6777 p. 272–281.
[9] Guo X, Xu S, Che W, Zhang X. Automatic Motion Generation Based on Path Editing from Motion Capture Data. In Lecture Notes in Computer Science. 2010; Volume 6250, p. 91-104.
[10] Tilmanne J, Dutoit T. Expressive gait synthesis using PCA and Gaussian modelling, in: Motion Games, Springer. 2010; p. 363–374.
[11] Maiocchi R. 3-D character animation using motion capture. In Interactive Computer Animation, N. Magnetat-Thalmann and D. Thalmann, Eds. Prentice-Hall, London. 1996; p. 10–39.
[12] Kovar L, Gleicher M. Automated Extraction and Parameterization of Motions in Large Data Sets. In: ACM SIGGRAPH. 2004.
[13] Min J, Chen Y, Chai J. Interactive Generation of Human Animation with Deformable Motion Models. In: ACM Transactions on Graphics. 2009.
[14] Wang X, Chen Q, Wang W. 3D human motion editing and synthesis: a survey. Compute Math Methods Med. 2014:104535.
[15] Grassia F. Practical parameterization of rotations using the exponential map. Journal of Graphics Tools. 1998; vol. 3, no. 3, p. 29–48.
[16] Shoemake K. Animating Rotations with Quaternion Curves. In Brian A. Barsky, editor, Computer Graphics (SIGGRAPH ’85 Proceedings). 1985; volume 19, pages 245-254.
[17] Johnson M. Exploiting Quaternions to Support Expressive Interactive Character Motion. PhD thesis, Massachusetts Institute of Technology. 2003.
[18] Min J, Chai J. Motion Graphs++: A Compact Generative Model for Semantic Motion Analysis and Synthesis. In ACM Transactions on Graphics. 2012; 31(6).
[19] Kovar L, Gleicher M. Flexible automatic motion blending with registration curves. In ACM SIGGRAPH/EUROGRAPH Symposium on Computer Animation. 2003; 214–224.
[20] Myers C, Rabiner L. A comparative study of several dynamic time-warping algorithms for connected word recognition. In The Bell System Technical Journal. 1981; 60(7), p. 1389-1409.
[21] Ramsay J, Silverman B. Functional Data Analysis. Springer; 2005.